

19 - Prática: Visão Computacional (III)

Higor Miller Grassi

Descrição da Atividade: Este card é um video sobre a revolução da visão computacional e o processo enfrentado para otimizar as máquinas e fazerem elas enxergar e identificar os objetos na imagem, porem sem entender o que está acontecendo em sua volta.

Apesar do enorme potencial da visão computacional, mesmo os hardwares mais poderosos enfrentam dificuldades devido à imensa quantidade de informações que precisam ser processadas. sendo importante entender que capturar fotos ou vídeos não é o mesmo que "ver", assim como ouvir sons não equivale a compreendê-los. na visão, o processo começa nos olhos, mas de fato ocorre no cérebro, onde as informações visuais são interpretadas e transformadas em conhecimento útil.

A visão computacional busca replicar esse complexo processo humano, sendo capaz de nomear objetos, identificar pessoas, perceber geometria em três dimensões, compreender ações, emoções e intenções, no entanto, ainda existem muitos desafios, como por exemplo, as máquinas muitas vezes são treinadas para reconhecer padrões específicos, mas esses padrões podem não se aplicar ao contexto real ou variar inesperadamente. Como citado em um exemplo do vídeo, ao ensinar uma máquina a reconhecer um gato, ela aprende características como formato, textura e outras informações associadas, mas, se o gato estiver em uma posição incomum, como de ponta-cabeça, o modelo pode não ser capaz de identificá-lo corretamente, resultando em erros.

As crianças humanas não precisam ser explicitamente ensinadas a "enxergar". já que desde cedo, elas aprendem por meio de experiências visuais acumuladas, processando e interpretando uma quantidade imensa de informações. Até os três anos de idade, por exemplo, uma criança já viu centenas de milhões de "imagens" do mundo à sua volta, adquirindo uma compreensão natural do ambiente em que vive, esse processo contínuo e adaptativo é algo que a visão computacional tenta replicar, mas ainda de forma limitada.

Portanto, para que a visão computacional seja eficaz, é essencial treiná-la com conjuntos de dados extensos e de alta qualidade, garantindo que os modelos consigam lidar com diferentes contextos e variações. Os algoritmos avançados usados na visão computacional moderna são inspirados no cérebro humano e sua rede de neurônios. Um exemplo citado no vídeo descreve um algoritmo com impressionantes 24 milhões de nós, 140 milhões de parâmetros e 15 bilhões de conexões, e apesar disso, esses sistemas apenas "veem" e não conseguem assimilar ou interpretar com precisão o que está acontecendo de maneira contextual, ainda cometendo erros consideráveis. No futuro, o avanço da visão computacional terá aplicações revolucionárias em diversas áreas, como a medicina, onde poderá auxiliar na detecção precoce de doenças e no planejamento de cirurgias, na ciência, permitindo análises de dados

em escala antes impossíveis e na indústria, otimizando processos de automação e controle de qualidade.

Embora ainda estejamos longe de alcançar a complexidade e a precisão do cérebro humano, o progresso na área é promissor, os modelos se tornam mais sofisticados e os dados utilizados para treiná-los se tornam mais ricos e variados, a visão computacional terá um impacto cada vez maior na sociedade, transformando o modo como interagimos com máquinas e como essas, por sua vez, entendem o mundo ao nosso redor.

Conclusão: A atividade destacou os avanços e desafios da visão computacional, evidenciando seu potencial revolucionário em diversas áreas, enquanto ainda busca alcançar a complexidade e adaptabilidade do cérebro humano, juntamente com as dificuldades encontradas.